PCT WORLD INTELLECTUAL PROPERTY ORGANIZATION
International Bureau

INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification 6 :
H04M 3/50, H04Q 1/46, G10L 3/00

A1

(11) International Publication Number: **WO 98/24225**

(43) International Publication Date: 4 June 1998 (04.06.98)

(21) International Application Number: PCT/GB97/03231

(22) International Filing Date: 26 November 1997 (26.11.97)

(30) Priority Data:
96308590.7 28 November 1996 (28.11.96) EP
(34) Countries for which the regional or international application was filed: AT et al.

(71) Applicant *(for all designated States except US)*: BRITISH TELECOMMUNICATIONS PUBLIC LIMITED COMPANY [GB/GB]; 81 Newgate Street, London EC1A 7AJ (GB).

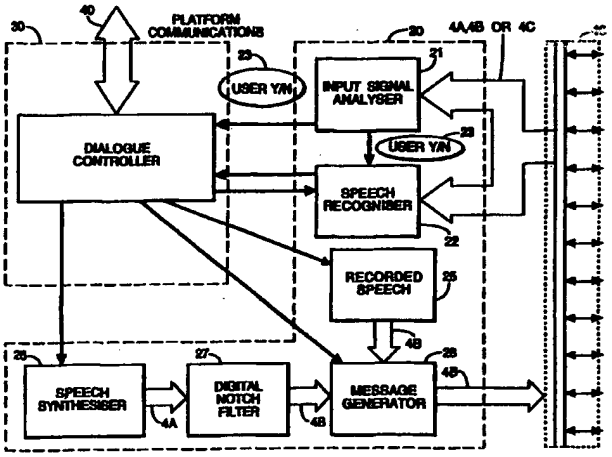
(72) Inventor; and
(75) Inventor/Applicant *(for US only)*: JOHNSTON, Robert, Denis [GB/GB]; 25 Haughgate Close, Woodbridge, Suffolk IP12 1LO (GB).

(74) Agent: LLOYD, Barry, George, William; BT Group Legal Services, Intellectual Property Dept., 8th floor, Holborn Centre, 120 Holborn, London EC1N 2TE (GB).

(81) Designated States: AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, CA, CH, CN, CU, CZ, DE, DK, EE, ES, FI, GB, GE, GH, HU, ID, IL, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MD, MG, MK, MN, MW, MX, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, UA, UG, US, UZ, VN, YU, ZW, ARIPO patent (GH, KE, LS, MW, SD, SZ, UG, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, ML, MR, NE, SN, TD, TG).

Published
With international search report.

(54) Title: INTERACTIVE APPARATUS



(57) Abstract

A problem exists with known interactive apparatuses, particularly those connected to a telecommunications network, in that it is difficult to distinguish between the echo of the outgoing prompt from the apparatus and the user's response to that prompt. An interactive apparatus is disclosed which allows the user to interrupt an outgoing prompt, and which removes a component which is norm ally found in the users' responses (e.g. a frequency band) from the outgoing prompt. An input signal analysis unit (21) in the apparatus is able to detect the response of the user by noting the presence of the component which is lacking from the outgoing prompt. As an alternative to removing the frequency band from the outgoing prompt, the apparatus may force spaced timeslots in the outgoing signal to silence. In that case, the input signal analysis unit can detect the presence of the user's response on determining that no periods of silence have been observed in the input signal over a pre–determined time interval. As well as being applicable to apparatuses which involve the user in prompt/response dialogues, the invention is also useful in relation to the interruption of messages being replayed by voice–controllable answerphones or the like.

FOR THE PURPOSES OF INFORMATION ONLY

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AL	Albania	**ES**	Spain	**LS**	Lesotho	**SI**	Slovenia
AM	Armenia	**FI**	Finland	**LT**	Lithuania	**SK**	Slovakia
AT	Austria	**FR**	France	**LU**	Luxembourg	**SN**	Senegal
AU	Australia	**GA**	Gabon	**LV**	Latvia	**SZ**	Swaziland
AZ	Azerbaijan	**GB**	United Kingdom	**MC**	Monaco	**TD**	Chad
BA	Bosnia and Herzegovina	**GE**	Georgia	**MD**	Republic of Moldova	**TG**	Togo
BB	Barbados	**GH**	Ghana	**MG**	Madagascar	**TJ**	Tajikistan
BE	Belgium	**GN**	Guinea	**MK**	The former Yugoslav Republic of Macedonia	**TM**	Turkmenistan
BF	Burkina Faso	**GR**	Greece			**TR**	Turkey
BG	Bulgaria	**HU**	Hungary	**ML**	Mali	**TT**	Trinidad and Tobago
BJ	Benin	**IE**	Ireland	**MN**	Mongolia	**UA**	Ukraine
BR	Brazil	**IL**	Israel	**MR**	Mauritania	**UG**	Uganda
BY	Belarus	**IS**	Iceland	**MW**	Malawi	**US**	United States of America
CA	Canada	**IT**	Italy	**MX**	Mexico	**UZ**	Uzbekistan
CF	Central African Republic	**JP**	Japan	**NE**	Niger	**VN**	Viet Nam
CG	Congo	**KE**	Kenya	**NL**	Netherlands	**YU**	Yugoslavia
CH	Switzerland	**KG**	Kyrgyzstan	**NO**	Norway	**ZW**	Zimbabwe
CI	Côte d'Ivoire	**KP**	Democratic People's Republic of Korea	**NZ**	New Zealand		
CM	Cameroon			**PL**	Poland		
CN	China	**KR**	Republic of Korea	**PT**	Portugal		
CU	Cuba	**KZ**	Kazakstan	**RO**	Romania		
CZ	Czech Republic	**LC**	Saint Lucia	**RU**	Russian Federation		
DE	Germany	**LI**	Liechtenstein	**SD**	Sudan		
DK	Denmark	**LK**	Sri Lanka	**SE**	Sweden		
EE	Estonia	**LR**	Liberia	**SG**	Singapore		

INTERACTIVE APPARATUS

The present invention relates to an interactive apparatus.

In recent years, an increasing number of everyday telephone interactions have been automated, thereby removing the need for a human operator to progress the interaction.

One of the first interactions to be automated was simply the leaving of a message for an intended recipient who was not present to take the call. Recently, more complex services such as telephone banking, directory enquiries and dial-up rail timetable enquiries have also been automated. Many answerphones now additionally offer a facility enabling their owner to telephone them and hear messages which have been left. Another service which has now been automated is the reading of stored e-mail messages over the telephone.

In each of the above cases, a user, in effect carries out a spoken dialogue with an apparatus which includes an interactive apparatus, the telephone he or she is using and elements of the Public Switched Telephone Network.

In the spoken dialogue it is often useful if the user is able to interrupt. For example, a user might wish to interrupt if he or she is able to anticipate what information is being requested part way through a prompt. The facility enabling interruption (known as a "barge-in" facility to those skilled in the art) is even more desirable in relation to message playback apparatuses (such as answerphones) where a user may wish to move onto another message without listening to intervening messages.

Providing a barge-in facility is made more difficult if some of the output from the interactive apparatus is fed-back to the input which receives the user's commands. This feedback arises owing to, for example, junctions in the network where voice-representing signals transmitted from the interactive apparatus are reflected back to its input. It is also caused by the acoustic echo of the speech output from the

speaker of the user's telephone back to the microphone (this is especially problematic in relation to handsfree operation). There is therefore a need to distinguish fed-back output signals from the user's input in order to provide a more reliable barge-in facility than has hitherto been possible.

According to the present invention there is provided an interactive apparatus comprising:

signal output means arranged in operation to output a signal representative of conditioned speech;

signal input means arranged in operation to receive a signal representative of a user's spoken command;

wherein the conditioned speech lacks a component normally present in speech;

command detection means operable to detect a user's command spoken during issuance of the conditioned speech by detecting the input of a signal which represents speech including the component lacking from the conditioned speech

The advantage of providing such an apparatus is that it is better able to detect the presence of a user's commands. This is particularly useful in relation to an apparatus which uses a conventional speech recogniser, as the performance of such recognisers falls off sharply if the voice signal they are analysing is in any way corrupted. In an interactive apparatus distortion caused by an echo of the interactive apparatus's output can cause the user's command to be corrupted. The present invention alleviates this problem by enabling the apparatus to stop outputting voice-representing signals or speech as soon as the user's response is detected.

In some embodiments, the apparatus further comprises a means for conditioning signals representative of speech output by the interactive apparatus. Because the quality of recorded speech is better than the quality of speech synthesised by conventional synthesisers, many conventional interactive apparatuses use recorded speech for those parts of the dialogue which are frequently used. However, for apparatuses such as those which are required to output signals representing a

spoken version of various telephone numbers or amounts of money it is currently impractical to record a spoken version of every possible output. Hence, such outputs are synthesised when required. A recorded speech signal can be pre-conditioned to lack the said component at the time that the speech signal is recorded. Hence, apparatuses whose entire output is recorded speech do not require a means for conditioning the signals representative of speech to be output by the interactive apparatus. Such apparatuses have the clear advantage of being less complex in their construction and are hence cheaper to manufacture.

Preferably, the said lacking component comprises one or more portions of the frequency spectrum. This has the advantage that the apparatus is easy to implement.

The apparatus is found to be most effective when the portion of the frequency spectrum lies in the range 1000 Hz to 1500 Hz.

Preferably, the width of the frequency band is in the range 80 Hz to 120 Hz. It is found that if the width of the frequency band is greater than 120 Hz then the output which the user hears is significantly corrupted, whereas if the width is less than 80 Hz the conditioning of the output of the interactive apparatus is made more difficult and it also becomes harder to discriminate between situations where the user is speaking and situations where he or she is not.

According to a second aspect of the present invention there is provided a method of detecting a user's spoken command to an interactive apparatus, said method comprising the steps of:

outputting a signal representative of conditioned speech, wherein the conditioned speech lacks a component normally comprised in users' spoken commands;

monitoring signals input to the interactive apparatus for the presence of signals representative of speech including said component; and

determining that the input signal represents the user's spoken command on detecting the presence of signals representative of speech including said component.

According to a third aspect of the present invention there is provided a voice-controllable apparatus comprising:

an interactive apparatus according to the first aspect of the present invention;

means for converting said signal representative of conditioned speech to conditioned speech; and

means for converting a user's spoken command to a signal representative thereof.

The problems addressed by the present invention also occur in relation to apparatuses which are directly voice-controlled (i.e. where there is no intermediate communications network). Embodiments of the third aspect of the present invention therefore include, amongst other things, domestic and work-related apparatuses such as personal computers, televisions, and video-recorders offering interactive voice control.

There now follows a detailed description of a specific embodiment of the present invention. This description is given by way of example only, with reference to the accompanying drawings, in which:

Figure 1 is a functional block diagram of part of an automated telephone banking apparatus installed in a communications network;

Figure 2 is a flow diagram representing the progress of a dialogue with a first time user of the apparatus;

Figure 3 is a diagram illustrating the progress of the same dialogue with a more experienced user;

Figure 4A illustrates the spectrum of the user's voice;

Figure 4B illustrates the spectrum of the signal output by the apparatus; and

Figure 4C illustrates the spectrum of the user's voice corrupted by an echo of the apparatus's output.

Figure 1 illustrates a signal processing unit used in providing an automated telephone banking service. In practice, the speech processing unit will be connected by an FDDI (Fibre Distributed Data Interface) local area network to a number of other units such as a telephone signalling unit, a file server unit for providing a large database facility, an assistant back up and data collection unit and an element management unit. A suitable apparatus for providing such a service is the interactive speech applications platform manufactured by Ericsson Ltd.

The speech processing unit (Figure 1) is interfaced to the telecommunications network via a digital line interface 10. The digital line interface inputs the digital signals which represents the user's voice from the telecommunications network and outputs this digital signal to the signal processing unit 20. The digital line interface 10 also inputs signals representing the spoken messages output by the apparatus from the signal processing unit 20 and modifies them to a form suitable for transmission over the telecommunications network before outputting those signals to the network. The digital line interface 10 is capable of handling a large number of incoming and outgoing signals simultaneously.

A signal processing unit 20 inputs the modified signals representing the user's voice from the digital line interface 10 and carries out a series of operations on those signals under the control of a dialogue controller 30 before outputting a signal representing the spoken response to the user via the digital line interface 10. The signal processing unit 20 includes four output processors 25, 26, 27, 28 and two input processors 21, 22.

The recorded speech output processor 25 is arranged to output a digital signal representing one of a number of messages stored therein which are frequently output by the apparatus. The particular message to be output is determined in accordance with a parameter supplied from the dialogue controller 30. The speech synthesiser processor 26 is used to output digital signals representing synthesised speech. The content of the spoken message is determined by the dialogue controller 30 which sends alphanumeric data representing the content of the message to the speech synthesiser processor 26.

The signal output by the speech synthesiser 26 is input to a digital notch filter 27. For reasons which will be explained below, this filter 27 is arranged to remove components of the synthesised signal lying in a frequency band from 1200 Hz to 1300 Hz. It will be realised that by those skilled in the art that although the speech synthesiser 26 and digital notch filter 27 are illustrated as separate processors, the two functions may be provided on a single processor.

The messages stored in the recorded speech processor 25 are recorded using a filter with a similar transfer function to the digital notch filter 27. Thus, the output of the speech synthesiser processor 26 might have a spectrum similar to that illustrated in Figure 4A, whereas the output of the digital notch filter 27 or the recorded speech processor 25 might have a spectrum similar to that shown by the solid line in Figure 4B.

The outputs of the filter 27 and the recorded speech processor 25 are passed to a message generator 28 which, for messages which have both a synthesised portion and a recorded speech portion, concatenates the two parts of the message before outputting the concatenated message via the digital line interface 10 to the user.

The two input signal processors are an input signal analyser 21 and a speech recogniser 22.

The input speech analyser 21 receives the signal representing the user's voice from the digital line interface 10 and passes it through a bandpass filter whose

passband extends from 1200 Hz to 1300 Hz. Thereafter, the input signal analyser compares the output of the bandpass filter with a threshold T (see Figure 4). If the signal strength in the passband lies above the threshold then the input signal analyser outputs a "user present" signal 23 indicative of the fact that the signal being input to it comprises the user's voice. On the other hand, if the signal strength within the passband falls below the threshold, then the analyser outputs an alternative version of the signal 23 to indicate that the signal input to the signal analyser 21 does not comprise the user's voice.

The incoming speech representing signal is also input to the speech recogniser 22 which is supplied with possible acceptable responses by the dialogue controller 30. On the user present signal 23 indicating that the user's voice is comprised in the input signal, the speech recogniser attempts to recognise the current word being spoken by the user and outputs the result to the dialogue controller 30.

The dialogue controller 30 then responds to the word or word spoken by the user in accordance with the software controlling it and controls the output processors in order to provide the user with a suitable response.

A dialogue (Figure 2) between the automated banking apparatus and an inexperienced user is initiated by the user dialling the telephone number of the apparatus. Once the user is connected to the apparatus the dialogue controller 30 instructs the recorded speech processor 25 to output a welcome message R1, immediately followed by an account number requesting prompt R2. As mentioned above, all recorded messages and prompts stored within the recorded speech processor 25 are recorded so as to have a spectrum similar to the one illustrated by the solid line in Figure 4B. Figure 4B shows that the spectrum of the recorded messages lacks any components having a frequency between 1200 Hz and 1300 Hz, but is otherwise normal. On outputting the message, it may be that an echo in the message is received back at the input signal processors 21, 22. Although it is likely that the spectrum will be altered slightly by the reflection process, the reflection process will not introduce frequencies which were not present in the outgoing signal and hence will not introduce frequencies in the frequency band

1200 Hz to 1300 Hz. Nevertheless, it is likely that some noise will be added to the output signal whilst it is being transmitted from the output signal processes 25, 26, 27, 28 to the input signal processes 21, 22. Hence, the spectrum of the echo may be similar to that shown as a dashed line in Figure 4B.

Returning to Figure 1, the echo of the prompt R2 is received at the input signal analyser 21 where it is bandpass filtered (the passband extending between 1200 Hz and 1300 Hz), and the resulting signal is compared to a threshold T. Since the echo of the outgoing prompt does not contain a significant component in the frequency band 1200 Hz to 1300 Hz, the signal falls below the threshold and the input signal analyser 21 outputs the signal 23 indicating, throughout the duration of the prompt R2, that the user is not speaking.

The user then proceeds to enter his account number using the DTMF (Dual Tone Multiple Frequency) keys on his phone. These tones are received by the speech recogniser 22 which converts the tones into numeric data and passes them to the dialogue controller 30. The dialogue controller 30 then forwards the account number to a customer database file server provided on the FDDI local area network. The file server then returns data indicating what services are to be made available in relation to this account and other data relating to the customer such as a personal identification number (PIN). Although not shown in Figures 2 and 3, the system will ask for the customer to enter his PIN immediately after having requested his account number.

The dialogue controller 30 then instructs the recorded speech processor 25 to output a type-of-service-required prompt R3 which the user listens to before replying by saying the word "transfer". The user's voice might have a spectrum similar to that shown in Figure 4A. When a signal representing his voice is passed to the input signal analyser 21, it is found that the signal contains a significant component from the frequency band 1200 Hz to 1300 Hz and hence the input to analyser 21 outputs a signal 23 indicative of the fact the user is speaking to the speech recogniser 22. The speech recogniser 22 recognises the word currently

being input to the apparatus to be "transfer" and passes a signal indicating that that is the word received to the dialogue controller 30.

As a result of having received this response, the dialogue controller 30 then instructs the recorded speech processor 25 to output a prompt asking the user how much money he wishes to transfer. The user then replies saying the amount of money he wishes to transfer, spoken entry of this information being potentially more reliable than information from the telephone keypad because a mistake in entering the DTMF tones may result in the user requesting the transfer of an amount of money which is an order of magnitude more or less than he would wish to transfer.

The user's response is then processed by the speech recogniser 22 and data indicating how much money the user has requested to transfer (£316.17 in this example) is passed the dialogue controller 30. The dialogue controller 30 then instructs the recorded speech processor 25 to send the recorded speech messages "I heard" and "is that correct?" to the message generator 28. The dialogue controller 30 then instructs the speech synthesiser 26 to synthesise a spoken version of £316.17. A synthesised version of these words is output by the speech synthesiser 26 and has a spectrum similar to that shown in Figure 4A. The signal is then passed through the digital notch filter 27 and is output having a spectrum similar to the solid line spectrum of Figure 4B. The modified synthesised message is then loaded into the message generator 28.

The message generator 28 then concatenates the two recorded speech messages and the synthesised speech message to provide the prompt R5 which is output via the digital line interface 10 to the user. The dialogue then continues.

A user who is more familiar with the system may carry out a dialogue like that shown in Figure 3. The initial part of the dialogue is identical to that described in relation to Figure 2 until the user interrupts the account number requesting prompt R2, using his telephone keypad to enter his account number. The DTMF tones output by his telephone are input to the speech recogniser 22 which converts the

tones to the account number representing the data and passes that data to the dialogue controller 30. As soon as the dialogue controller 30 receives this data it sends a signal to the recorded speech processor 25 to halt the output of the account number requesting prompt R2. Clearly, once the apparatus has stopped issuing the prompt R2, no echo of that prompt will be received back at the apparatus. Hence, the speech recogniser can recognise the other DTMF tones input by the user without the presence of the interfering echo.

The dialogue then continues as before until the user interrupts the service required prompt R3 by saying the word "transfer". During the first two words of the message R3, it will be realised that the input signal analyser 21 will be outputting a signal 23 which indicates that the user's voice is not present. However, as the user interrupts the output message, the signal received at the apparatus will be a combination of the user's voice and an echo of the outgoing prompt. The spectrum of this combination signal will be similar to that of the user's voice alone (Figure 4A), but because the spectrum of the echo signal lacks any components between 1200 Hz and 1300 Hz, will feature a small notch between 1200 Hz and 1300 Hz. (Figure 4C).

The combination signal is passed to the input signal analyser 21 where it is passed through a bandpass filter and found to have a significant component in the frequency range 1200 Hz to 1300 Hz. The input signal analyser 21 therefore outputs a signal 23 (indicating that the user's voice is present) to both the speech recogniser 22 and the dialogue controller 23. On receiving the signal 23, the dialogue controller 30 instructs the recorded speech processor 25 to halt its output of the prompt R3. Soon after, the echo of the prompt ceases to be a component for signals received at the speech recogniser 22, and the recogniser is better able to recognise the word currently being spoken by the user. Once the response of the user has been recognised, it is passed to the dialogue controller 30.

Thereafter, the user interrupts the next two prompts of the dialogue in a similar way to the way in which he interrupted the type-of-service-required prompt R3.

It will be realised that in the above embodiment, the component lacking from the pre-conditioned spoken prompt comprises a portion of the frequency spectrum. However, it is also envisaged that other components might be lacking. For example, timeslots of short duration (say 1 to 5ms) could be removed from the spoken prompt at a regular interval (say every 20ms to 100ms). If, for example, the speech is digitally sampled at 8kHz, this might be achieved by setting 8 to 40 samples to a zero value at an 160-800 sample interval. To take a particular value, if 20 samples were to be removed from the signal at a 400 sample interval, then the input signal analyser might be set up such that if it did not detect a corresponding silence or near silence (i.e. where the volume is below a given threshold) during a received signal duration of 800 samples, then it might output a signal indicative that the user is speaking.

It will be seen how the "barge-in" facility allows the user to carry out his transaction more quickly. More importantly, by being able to interrupt the prompt issued by the apparatus in this way, the user feels more in control of the dialogue.

CLAIMS

1. An interactive apparatus comprising:

signal output means arranged in operation to output a signal representative of conditioned speech;

signal input means arranged in operation to receive a signal representative of a user's spoken command;

wherein the conditioned speech lacks a component normally present in speech;

command detection means operable to detect a user's command spoken during issuance of the conditioned speech by detecting the input of a signal which represents speech including the component lacking from the conditioned speech.

2. An apparatus according to claim 1 which further comprises a means for conditioning a signal representing speech so as to provide said signal representative of conditioned speech.

3. An apparatus according to Claim 2 wherein said conditioning means comprises a digital filter.

4. An apparatus according to any preceding claim in which the lacking component comprises one or more portions of the frequency spectrum.

5. An apparatus according to Claim 4 wherein the mid-point of said portion lies in the range 1000 Hz to 1500 Hz.

6. An apparatus according to Claim 5 in which the mid-point lies in the range 1200 Hz to 1300 Hz.

7. An apparatus according to any one of Claims 4 to 6 in which the width of said portion falls within the range 80 Hz to 120 Hz.

8. An apparatus according to any one of Claims 1 to 3 wherein the lacking component comprises a plurality of spaced short time-segments of said speech signal.

9. A voice-controllable apparatus comprising:

an interactive apparatus according to any preceding claim;

means for converting said signal representative of conditioned speech to conditioned speech; and

means for converting a user's spoken command to a signal representative thereof.

10. A method of detecting a user's spoken command to an interactive apparatus, said method comprising the steps of:

outputting a signal representative of conditioned speech, wherein the conditioned speech lacks a component normally comprised in users' spoken commands;

monitoring signals input to the interactive apparatus for the presence of signals representative of speech including said component; and

determining that the input signal represents the user's spoken command on detecting the presence of signals representative of speech including said component.

11. A method according to Claim 10 further comprising the step of conditioning the signal representative of the spoken command.

12. An apparatus substantially as hereinbefore described with reference to and as illustrated in the accompanying drawings.

13. A method of detecting a user's response to a prompt issued by an interactive apparatus substantially as hereinbefore described with reference to and as illustrated in the accompanying drawings.

14. A communications network including an apparatus according to any one of Claims 1 to 8.

15. An interactive apparatus comprising:

output means arranged in operation to output a pre-conditioned spoken prompt or a signal representative thereof;

input means arranged in operation to input a signal representative of a user's voice;

wherein the pre-conditioned spoken prompt lacks a component normally present in speech;

response detection means operable, during the issuance of the pre-conditioned prompt, to detect any input from the user by detecting the input of a signal which comprises the component lacking from the prompt.

Fig.1.

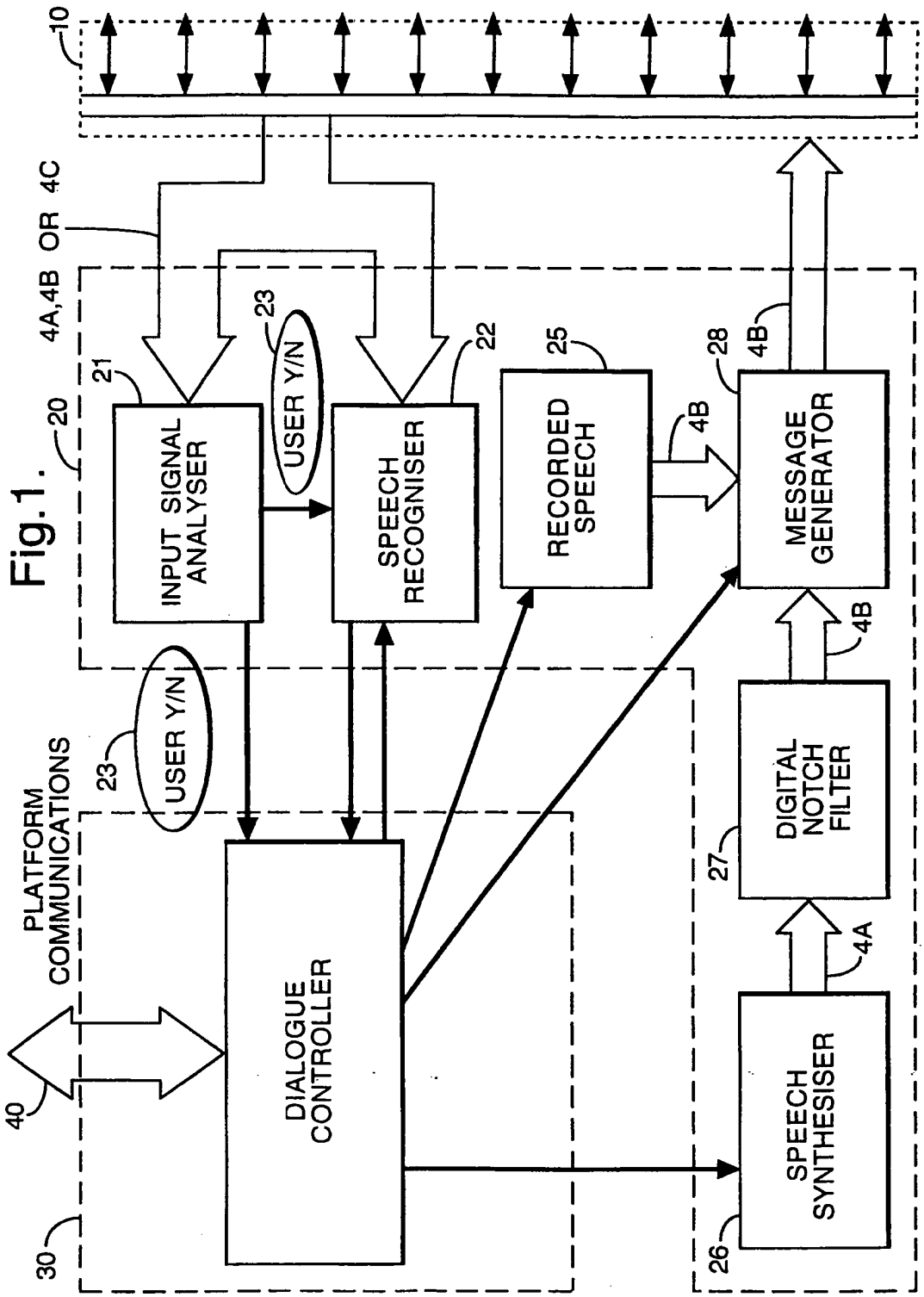

10
4C OR
4A,4B
20
21
INPUT SIGNAL ANALYSER
23
USER Y/N
22
SPEECH RECOGNISER
25
RECORDED SPEECH
28
4B
MESSAGE GENERATOR
4B
DIGITAL NOTCH FILTER
27
4A
SPEECH SYNTHESISER
26
PLATFORM COMMUNICATIONS
23
USER Y/N
DIALOGUE CONTROLLER
40
30

Fig.2.

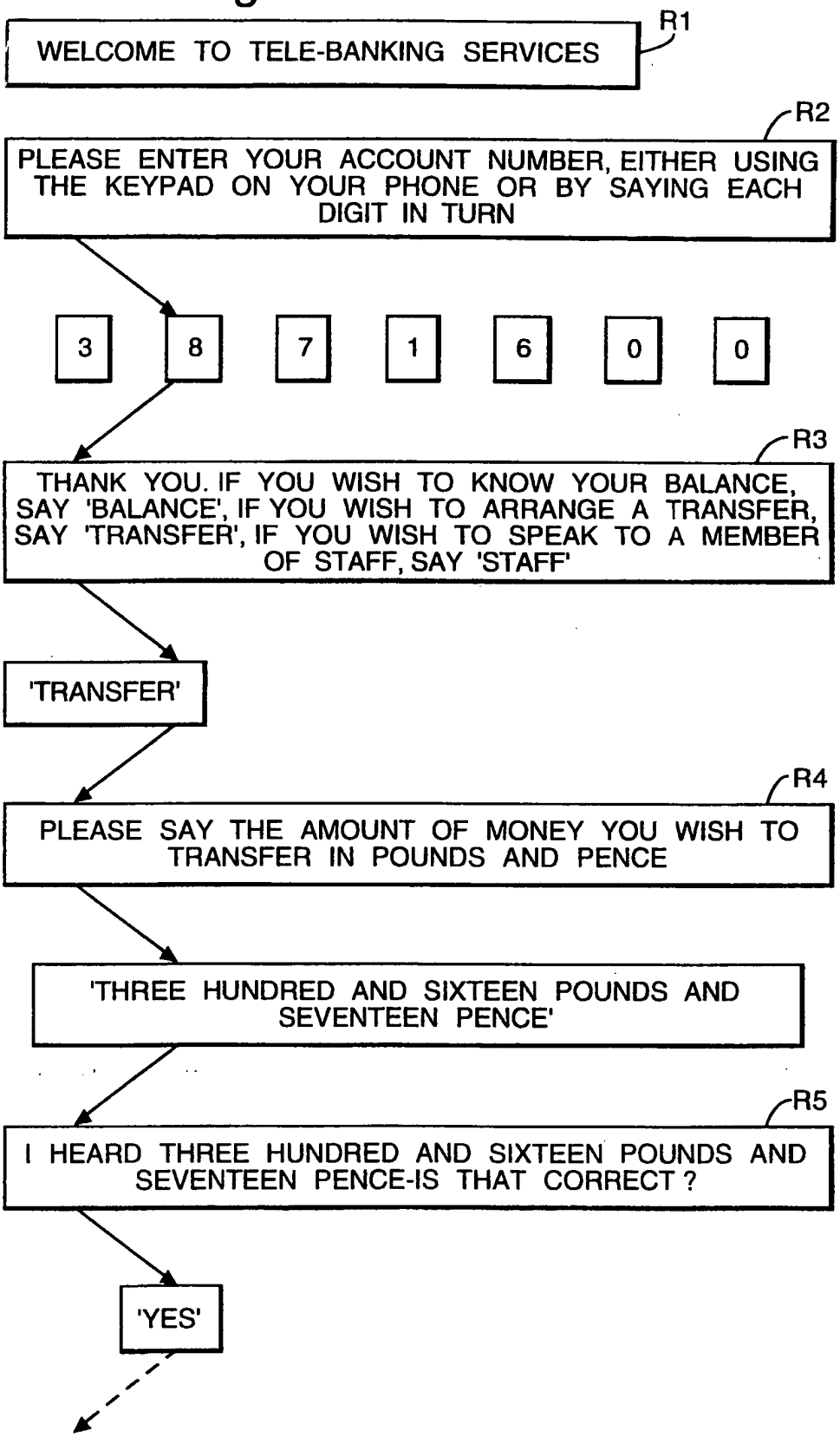

R1
WELCOME TO TELE-BANKING SERVICES
R2
PLEASE ENTER YOUR ACCOUNT NUMBER, EITHER USING THE KEYPAD ON YOUR PHONE OR BY SAYING EACH DIGIT IN TURN
3
8
7
1
6
0
0
R3
THANK YOU. IF YOU WISH TO KNOW YOUR BALANCE, SAY 'BALANCE', IF YOU WISH TO ARRANGE A TRANSFER, SAY 'TRANSFER', IF YOU WISH TO SPEAK TO A MEMBER OF STAFF, SAY 'STAFF'
'TRANSFER'
R4
PLEASE SAY THE AMOUNT OF MONEY YOU WISH TO TRANSFER IN POUNDS AND PENCE
'THREE HUNDRED AND SIXTEEN POUNDS AND SEVENTEEN PENCE'
R5
I HEARD THREE HUNDRED AND SIXTEEN POUNDS AND SEVENTEEN PENCE-IS THAT CORRECT ?
'YES'

Fig.3.

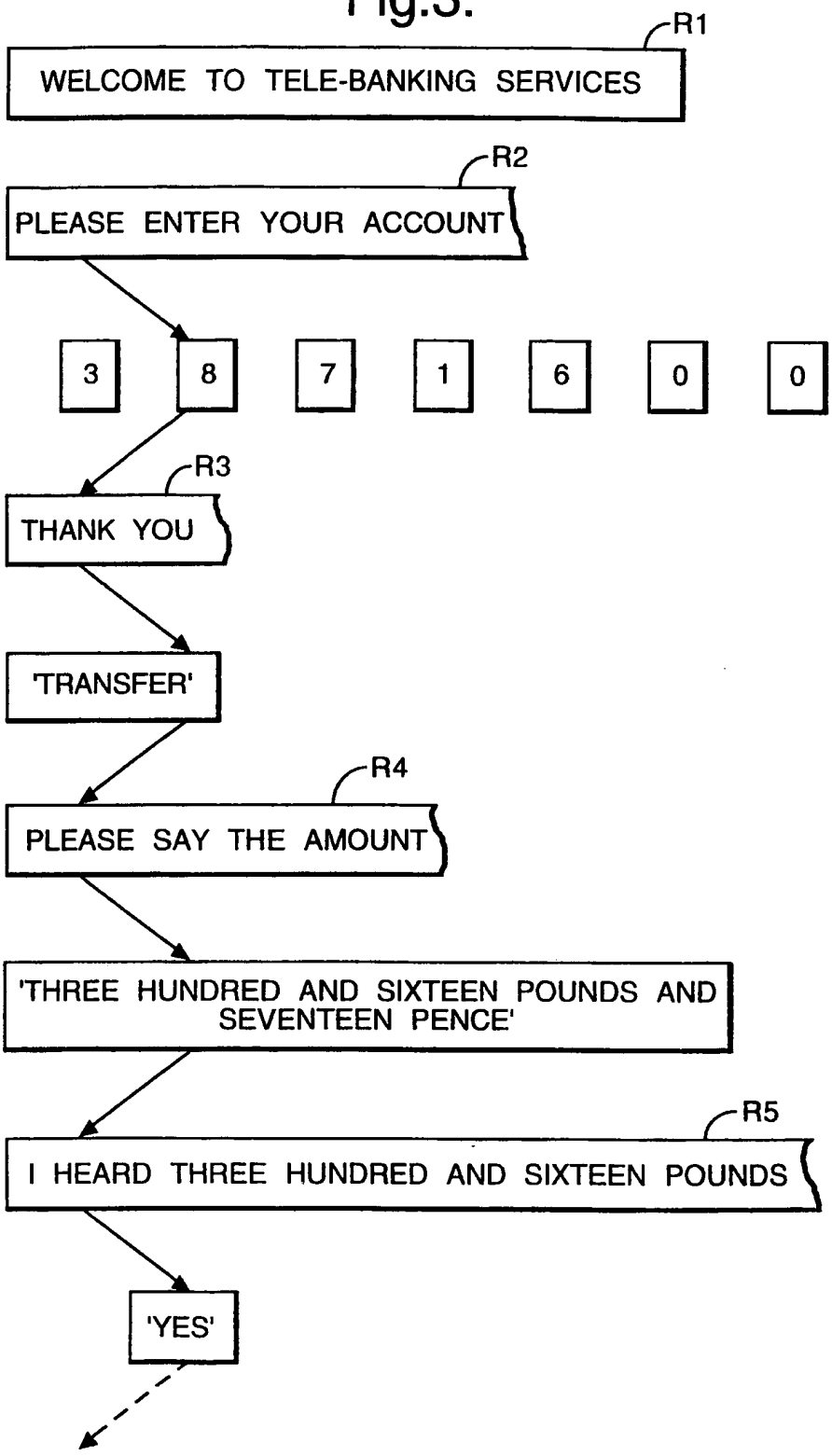

R1
WELCOME TO TELE-BANKING SERVICES
R2
PLEASE ENTER YOUR ACCOUNT
3
8
7
1
6
0
0
R3
THANK YOU
'TRANSFER'
R4
PLEASE SAY THE AMOUNT
'THREE HUNDRED AND SIXTEEN POUNDS AND SEVENTEEN PENCE'
R5
I HEARD THREE HUNDRED AND SIXTEEN POUNDS
'YES'

Fig.4A.

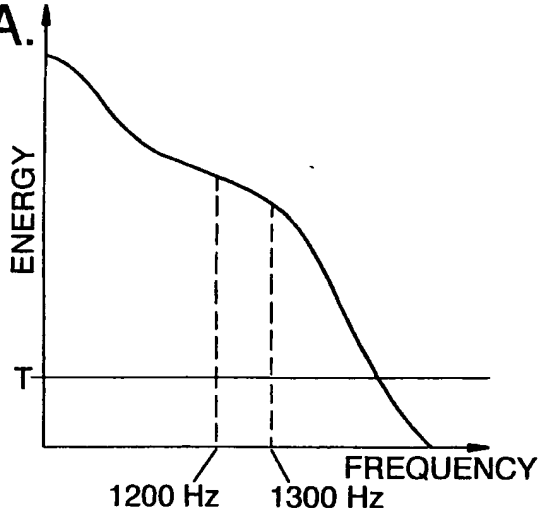


Fig.4B.

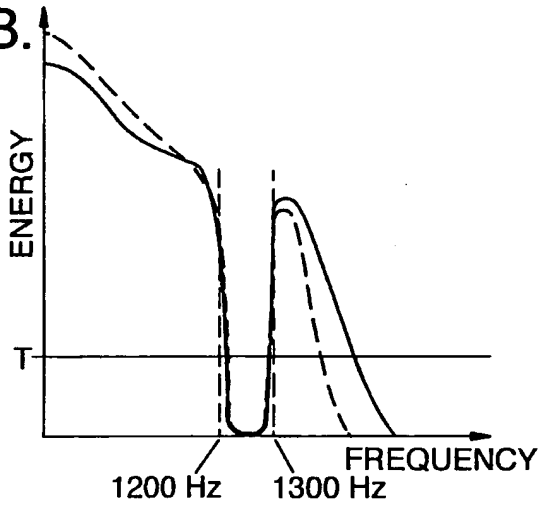
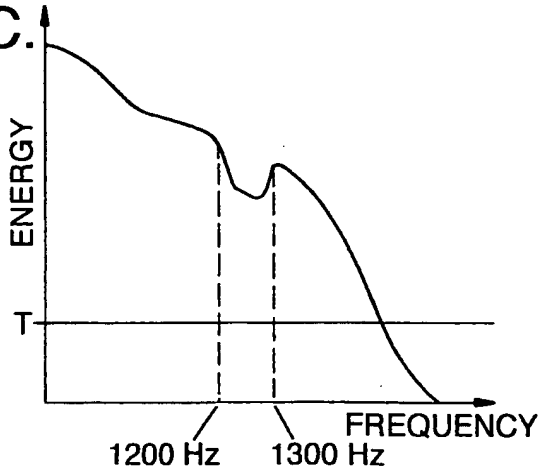


Fig.4C.



INTERNATIONAL SEARCH REPORT

In ational Application No
PCT/GB 97/03231

A. CLASSIFICATION OF SUBJECT MATTER
IPC 6 H04M3/50 H04Q1/46 G10L3/00

According to International Patent Classification(IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)
IPC 6 H04M H04Q G10L

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practical, search terms used)

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category °	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	US 4 191 855 A (SAKAI HISAO) 4 March 1980 see the whole document ---	1-4, 10-15
X	US 4 521 647 A (OLSON PETER D ET AL) 4 June 1985 see column 4, line 54 - column 5, line 25 ---	1-4, 10-15
A	US 4 914 692 A (HARTWELL WALTER T ET AL) 3 April 1990 see abstract ---	
A	GB 2 251 765 A (TELSIS LIMITED) 15 July 1992 -----	

☐ Further documents are listed in the continuation of box C.

☒ Patent family members are listed in annex.

° Special categories of cited documents :

"A" document defining the general state of the art which is not considered to be of particular relevance

"E" earlier document but published on or after the international filing date

"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)

"O" document referring to an oral disclosure, use, exhibition or other means

"P" document published prior to the international filing date but later than the priority date claimed

"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone

"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art.

"&" document member of the same patent family

1

Date of the actual completion of the international search	Date of mailing of the international search report
3 March 1998	11/03/1998

Name and mailing address of the ISA	Authorized officer
European Patent Office, P.B. 5818 Patentlaan 2 NL - 2280 HV Rijswijk Tel. (+31-70) 340-2040, Tx. 31 651 epo nl, Fax: (+31-70) 340-3016	Vandevenne, M

Form PCT/ISA/210 (second sheet) (July 1992)

INTERNATIONAL SEARCH REPORT

Information on patent family members

Int tional Application No
PCT/GB 97/03231

Patent document cited in search report	Publication date	Patent family member(s)	Publication date
US 4191855 A	04-03-80	JP 53027332 A	14-03-78
		AU 510583 B	03-07-80
		AU 2822777 A	01-03-79
US 4521647 A	04-06-85	EP 0153125 A	28-08-85
		JP 60249459 A	10-12-85
US 4914692 A	03-04-90	CA 1326075 A	11-01-94
GB 2251765 A	15-07-92	NONE	